Extended Features using Machine Learning Techniques for Photo Liking Prediction

Steve Göring, Konstantin Brand, Alexander Raake

Audiovisual Technology Group, Technische Universität Ilmenau, Germany;
Email: [steve.goering, konstantin.brand, alexander.raake]@tu-ilmenau.de

May 31, 2018

TECHNISCHE UNIVERSITÄT ILMENAU

Motivation



- number of posted photos increases daily[1]
- hard for user to select or find good photos based on appeal or liking
- how to decide if a photo is good or not for a social media platform?
- is it possible to automatically predict if an image will be liked or not?

→ connection between human rating, aesthetic appeal and liking decisions

[1] for Flickr: average 1.68 million photos per day for 2016, see [fli18]

Human rating, aesthetic appeal and liking

TECHNISCHE UNIVERSITÄT ILMENAU

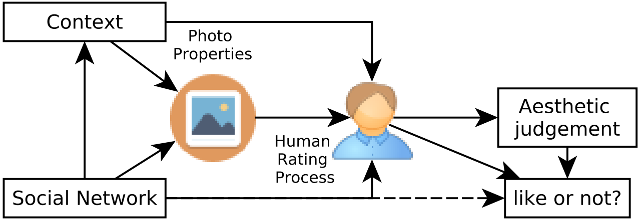


Figure: aesthetic and liking, based on Leder et al. [Led+04]

- aesthetics: artwork, intent of artist, genre of art, perception, . . . [Jos+11]
- main factors for human aesthetics judgment: the photo, the context and the social influences [Led+04]
- lines between aesthetics and liking get blurred [LRB16b; LRB16a]

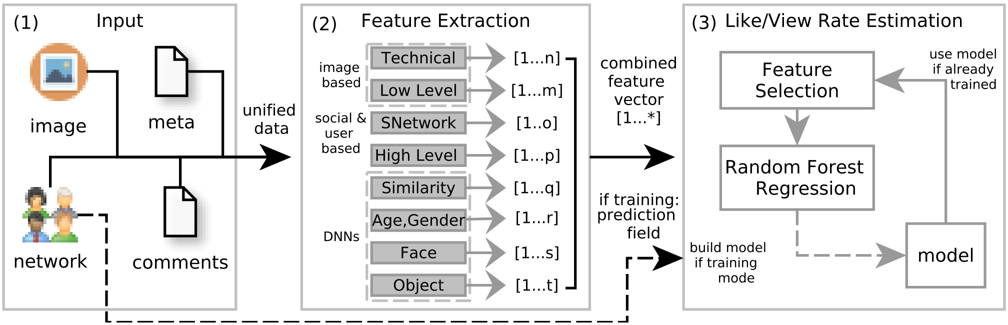
→ features and extensions

State of the art features and extensions

- many other works for photo aesthetic prediction published, e.g. [Dat+06; WBT10; KDH14; Kal+15; Tot+14]
- mostly only focused on classification problem
  - → focus on a regression problem formulation for liking prediction
- or using only low level features [Dat+06; WBT10], or titles and other data [KDH14]
  - → extend and combine feature ideas
  - → image, social network, natural language analysis and deep learning

→ our prediction framework

Our general approach



- features: technical, low-level (LLF), high-level, DNNs, social network (SN)
- focus on $log(\#like)/log(\#view)$ = like/view rate

→ example feature group: DNNs

Example feature group – DNN

TECHNISCHE UNIVERSITÄT ILMENAU

- using several **pre-trained** Deep Neural Networks (DNNs)
- **lastLayerValues**: inception V3 [Sze+15], similar approach for NR video quality [GSR18]
- **hash, hashProbs**: using image similarity DNN for image retrieval [Lin+15]
- **age, gender**: using DNN [RTV15; RTV16] for age and gender estimation
- **faceCount, maxProbFace**: DNN for face detection [PVZ15]

→ dataset for evaluation

Dataset – crawled images from 500px

Crawled 500px images and used samples for experiments.

category / sample	# crawled pictures	$log(\#like)/log(\#view)$ rate r min r	max r	$\bar{r}$	$\sigma(r)$
editors	20975	0.43	0.83	0.68	0.05
fresh	59130	0.10	0.86	0.62	0.11
sample9000	9000	0.10	0.82	0.62	0.11
combined20k	20000	0.11	0.83	0.65	0.09
all80k	80105	0.10	0.86	0.63	0.10

each sample shows similar distribution for $log(\#like)/log(\#view)$ rate

$\rightarrow$ evaluation

Evaluation of Feature Groups

Which feature groups are important? using *sample9000*

- *RMSE* and R^2 for leave-one-out experiment; $* =$ feature set

leave-out	technical	low-level	high-level	SN	DNNs	ALL*	LLF*
RMSE	0.098	0.098	0.098	0.104	0.097	0.098	0.108
R^2	0.233	0.228	0.229	0.135	0.255	0.231	0.072

- social network (SN) features show high impact
- better performance without DNNs (could be sample9000 property)
- single feature: show small impact (extended feature importance analysis)

$\rightarrow$ large scale evaluation

Evaluation of Photo-subject Derivable Features

TECHNISCHE UNIVERSITÄT ILMENAU

Can we just use photo derivable features?

- OPD= technical, LLF, DNN features; ALL, LLF= low level features

RMSE	ALL	OPD	LLF	R^2	ALL	OPD	LLF
sample9000	**0.098***	0.105	0.108		**0.231***	0.128	0.072
combined20k	**0.077***	0.088	0.091		**0.326***	0.114	0.056
all80k	**0.085***	0.096	0.099		**0.329***	0.139	0.091

- ALL performs best for all datasets
- OPD can be used to approximate image liking; better than LLF

$\rightarrow$ conclusion

Conclusion and Future Work

TECHNISCHE UNIVERSITÄT ILMENAU

- introduced a framework for image-liking prediction
  - pure image features; social media, meta information and comments
  - combination of image, social network, natural language analysis and deep learning
- performed evaluation of our new features
  - good results in comparision with only low-level features
  - large impact of social network features
  - still hard to predict liking
- open points, possible extensions:
  - other effects/indicators: temporal, psychological, fashion, art, . . .
  - modern aesthetic datasets

Thank you for your attention

. . . . . . are there any questions?

References I

[Dat+06] Ritendra Datta et al. "Studying aesthetics in photographic images using a computational approach". In: *European Conference on Computer Vision*. Springer. 2006, pp. 288–301.

[fli18] flickr. *Flickr Statistics*. 2018. URL: https://www.flickr.com/photos/franckmichel/6855169886/ (visited on 05/28/2018).

[GSR18] Steve Göring, Janto Skowronek, and Alexander Raake. "DeViQ – A deep no reference video quality model". In: *Electronic Imaging* (2018).

[Jos+11] D. Joshi et al. "Aesthetics and Emotions in Images". In: *IEEE Signal Processing Magazine* 28.5 (Sept. 2011), pp. 94–115.

References II

[Kal+15] Mahdi M Kalayeh et al. "How to Take a Good Selfie?" In: *Proceedings of the 23rd ACM international conference on Multimedia*. ACM. 2015, pp. 923–926.

[KDH14] Aditya Khosla, Atish Das Sarma, and Raffay Hamid. "What makes an image popular?" In: *Proceedings of the 23rd international conference on World wide web*. ACM. 2014, pp. 867–876.

[Led+04] Helmut Leder et al. "A model of aesthetic appreciation and aesthetic judgments". In: *British journal of psychology* 95.4 (2004), pp. 489–508.

[Lin+15] Kevin Lin et al. "Deep learning of binary hash codes for fast image retrieval". In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2015, pp. 27–35.

References III

[LRB16a] Pierre Lebreton, Alexander Raake, and Marcus Barkowsky. "Evaluation of aesthetic appeal with regard of user's knowledge". In: *EI* 2016.16 (2016), pp. 1–6.

[LRB16b] Pierre Lebreton, Alexander Raake, and Marcus Barkowsky. "Studying user agreement on aesthetic appeal ratings and its relation with technical knowledge". In: *QoMEX 2016*. IEEE. 2016, pp. 1–6.

[PVZ15] Omkar M Parkhi, Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition". In: *BMVC*. Vol. 1. 3. 2015, p. 6.

[RTV15] Rasmus Rothe, Radu Timofte, and Luc Van Gool. "DEX: Deep EXpectation of apparent age from a single image". In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2015, pp. 10–15.

References IV

TECHNISCHE UNIVERSITÄT ILMENAU

[RTV16] Rasmus Rothe, Radu Timofte, and Luc Van Gool. "Some like it hot-visual guidance for preference prediction". In: *arXiv preprint arXiv:1510.07867* (2016).

[Sze+15] Christian Szegedy et al. "Rethinking the Inception Architecture for Computer Vision". In: *CoRR* abs/1512.00567 (2015).

[Tot+14] Luam Catao Totti et al. "The impact of visual attributes on online image diffusion". In: *WebSci '14*. ACM. 2014, pp. 42–51.

[WBT10] Yaowen Wu, Christian Bauckhage, and Christian Thurau. "The good, the bad, and the ugly: Predicting aesthetic image labels". In: *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE. 2010, pp. 1586–1589.

Full list of features

TECHNISCHE UNIVERSITÄT ILMENAU

Feature groups, **S**=string, **I**=integer, **F**=float, **M**=multiple, SN=social network, LN=local network, * input dependent

Group	Name	Type	Src	Dim
Technical	cameraType, locationName	S	meta	2
	focalLength, ISOValue, shutterSpeed	I	meta	3
	latitude, longitude, aperture	F	meta	3
	height, width, dateInfos	I	meta	4
Low-Level	globalHue/Sat/Val	F	img	3
	subimageHue/Sat/Val 1...9	MF	img	27
	max/min index of SubImgHue/Sat/Val	MI	img	8
	colorDist, noiseDiff, edgeRatio	MF	img	*+2
High-Level	titleWordCount, -NonStopWordCount	I	meta	2
	titleWordLenDist	MI	meta	*
	titleTagJaccSim, -TagWordSim, -Sent	F	meta	3
DNNs	classDistScores, lastLayerValues	MF	img	2053
	hash, hashProbs	I	img	49
	age, gender, faceCount, maxProbFace	F	img	4
SN	commentSentiment, -WordLengthDist	MF	com	6+*
	comment, friendCommentRate	MF	com	*+1
	followers/friends/galleries/groups-Count	I	user	3
	userAffection/photos/favorites-Count	I	user	3
	LN-triangleCount/MeanFoFCount	I	user	2
	LN-2hopReachableUsers	I	user	1